PATENT

UNITED STATES APPLICATION FOR LETTERS PATENT

for

4 POINT DERATING SCHEME FOR PROPAGATION DELAY AND SETUP/HOLD TIME COMPUTATION

By

Sandeep Bhutani
14 Airway Cir. #3C
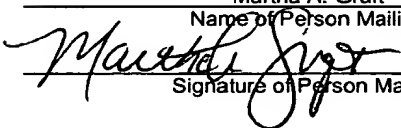Towson, MD 21286
Citizenship: India

Subramanian Venkateswaran
992 Henderson Avenue, #1
Sunnyvale, California 94086
Citizenship: India

4 POINT DERATING SCHEME FOR PROPAGATION DELAY SETUP/HOLD TIME COMPUTATION

BACKGROUND OF THE INVENTION

5   ### Field of the Invention

The present invention generally relates to the art of microelectronic integrated circuits. In particular, the present invention relates to the art of computing delays for cells in ASICs.

10   ### Description of the Prior Art

An integrated circuit chip (hereafter referred to as an "IC" or a "chip") comprises cells and connections between the cells formed on a surface of a semiconductor substrate. The IC may include a large number of cells and require complex connections between the cells.

15   A *cell* is a group of one or more circuit elements such as transistors, capacitors, and other basic circuit elements grouped to perform a function. Each of the cells of an IC may have one or more pins, each of which, in turn, may be connected to one or more other pins of the IC by wires. The wires connecting the pins of the IC are also formed on the surface of the chip.

20   A *net* is a set of two or more pins which must be connected. Because a typical chip has thousands, tens of thousands, or hundreds of thousands of pins which must be connected in various combinations, the chip also includes definitions of thousands, tens of thousands, or hundreds of thousands of nets, or sets of pins. All the pins of a net must be connected. The number of the nets for a chip is typically
25   in the same order as the order of the number of cells on that chip. Commonly, a majority of the nets include only two pins to be connected; however, many nets comprise three or more pins. Some nets may include hundreds of pins to be connected. A *netlist* is a list of nets for a chip.

Microelectronic integrated circuits consist of a large number of
30   electronic components that are fabricated by layering several different materials on a silicon base or wafer. The design of an integrated circuit transforms a circuit

1

description into a geometric description which is known as a *layout*. A layout consists of a set of planar geometric shapes in several layers.

The layout is then checked to ensure that it meets all of the design requirements. The result is a set of design files in a particular unambiguous

5 representation known as an intermediate form that describes the layout. The design files are then converted into pattern generator files that are used to produce patterns called masks by an optical or electron beam pattern generator.

During fabrication, these masks are used to pattern a silicon wafer using a sequence of photolithographic steps. The component formation requires

10 very exacting details about geometric patterns and separation between them. The process of converting the specifications of an electrical circuit into a layout is called the physical design.

Currently, the minimum geometric feature size of a component is on the order of 0.2 microns. However, it is expected that the feature size can be reduced

15 to 0.1 micron within the next few years. This small feature size allows fabrication of as many as 4.5 million transistors or 1 million gates of logic on a 25 millimeter by 25 millimeter chip. This trend is expected to continue, with even smaller feature geometries and more circuit elements on an integrated circuit, and of course, larger die (or chip) sizes will allow far greater numbers of circuit elements.

20 Due to the large number of components and the exacting details required by the fabrication process, physical design is not practical without the aid of computers. As a result, most phases of physical design extensively use Computer Aided Design (CAD) tools, and many phases have already been partially or fully automated. Automation of the physical design process has increased the level of

25 integration, reduced turn around time and enhanced chip performance.

The objective of physical design is to determine an optimal arrangement of devices in a plane or in a three dimensional space, and an efficient interconnection or routing scheme between the devices to obtain the desired functionality.

30

2

A. IC Configuration.

An exemplary integrated circuit chip is illustrated in Fig. 1 and generally designated by the reference numeral 26. The circuit 26 includes a semiconductor substrate 26A on which are formed a number of functional circuit blocks that can

5   have different sizes and shapes. Some are relatively large, such as a central processing unit (CPU) 27, a read-only memory (ROM) 28, a clock/timing unit 29, one or more random access memories (RAM) 30 and an input/output (I/O) interface unit 31. These blocks, commonly known as macroblocks, can be considered as modules for use in various circuit designs, and are represented as standard designs in circuit

10 libraries.

The integrated circuit 26 further comprises a large number, which can be tens of thousands, hundreds of thousands or even millions or more of small cells 32. Each cell 32 represents a single logic element, such as a gate, or several logic elements interconnected in a standardized manner to perform a specific function.

15 Cells that consist of two or more interconnected gates or logic elements are also available as standard modules in circuit libraries.

The cells 32 and the other elements of the circuit 26 described above are interconnected or routed in accordance with the logical design of the circuit to provide the desired functionality. Although not visible in the drawing, the various

20 elements of the circuit 26 are interconnected by electrically conductive lines or traces that are routed, for example, through vertical channels 33 and horizontal channels 34 that run between the cells 32.

B. Layout Design Process.

25 The input to the physical design problem is a circuit diagram, and the output is the layout of the circuit. This is accomplished in several stages including partitioning, floor planning, placement, routing and compaction.

1. Partitioning.

30 A chip may contain several million transistors. Layout of the entire circuit cannot be handled due to the limitation of memory space as well as the computation power available. Therefore, the layout is normally partitioned by

3

grouping the components into blocks such as subcircuits and modules. The actual partitioning process considers many factors such as the size of the blocks, number of blocks and number of interconnections between the blocks.

The output of partitioning is a set of blocks, along with the
5   interconnections required between blocks. The set of interconnections required is the netlist. In large circuits, the partitioning process is often hierarchical, although non-hierarchical (e.g. flat) processes can be used, and at the topmost level a circuit can have between 5 to 25 blocks. However, greater numbers of blocks are possible and contemplated. Each block is then partitioned recursively into smaller blocks.
10

2.   Floor planning and placement.

This step is concerned with selecting good layout alternatives for each block of the entire chip, as well as between blocks and to the edges. Floor planning is a critical step as it sets up the ground work for a good layout. During placement,
15   the blocks are exactly positioned on the chip. The goal of placement is to find a minimum area arrangement for the blocks that allows completion of interconnections between the blocks. Placement is typically done in two phases. In the first phase, an initial placement is created. In the second phase, the initial placement is evaluated and iterative improvements are made until the layout has minimum area
20   and conforms to design specifications.

3.   Routing.

The objective of the routing phase is to complete the interconnections between blocks according to the specified netlist. First, the space not occupied by
25   blocks, which is called the routing space, is partitioned into rectangular regions called channels. The goal of a router is to complete all circuit connections using the shortest possible wire length and using only the channel.

Routing is usually done in two phases referred to as the global routing and detailed routing phases. In global routing, connections are completed between
30   the proper blocks of the circuit disregarding the exact geometric details of each wire and terminal. For each wire, a global router finds a list of channels that are to be

4

used as a passageway for that wire. In other words, global routing specifies the loose route of a wire through different regions of the routing space.

Global routing is followed by detailed routing which completes point-to-point connections between terminals on the blocks. Loose routing is
5   converted into exact routing by specifying the geometric information such as width of wires and their layer assignments. Detailed routing includes the exact channel routing of wires.

In order for circuit designers to calculate the performance of ASICs, the designers need to compute the delays of the cells in the ASICs. In the present
10   invention, two types of delays are considered. The first type of delay is the propagation delay of a cell. A propagation delay of a cell is defined as the time duration a signal takes to travel from the input to the output of a cell. The measurement point at the input is called the switching threshold. The measurement point at the output is usually the 0.5 * Vdd (the power supply). A propagation delay
15   of a cell is defined for every input to output pin combination of a cell under both the rising and falling input conditions. The propagation delay is also affected by a given process (P), voltage (V) and temperature (T).

The second type of delay is the setup/hold time delay which is an input constraint for sequential cells. The setup time is defined as the time duration a data
20   signal is required to be available at the input of a cell before the clock signal transition, and the hold time is defined as the time duration a data signal is required to be stable after the clock signal transition. For the purpose of explanation, both propagation delay and setup/hold time, henceforth, will be referred as 'delay'.

The following derating equation is widely used in the industry to compute the
25   delay of a cell for a given P, V and T of a cell:

$D_{case} = K * D_{nom}$ where,

D_{nom} = nominal delay at nominal P, V and T (e.g., P = nominal process, T = room temperature, V = supply voltage);
30   D_{case} = delay for a given P, V and T;
$K = (1 + Kp) * ((1 + Kv (V_{case} - V_{nom})) * ((1 + Kt (T_{case} - T_{nom}))$;
$Kp = (D_{case} - D_{nom})/D_{nom}$;

5

$$Kv = ((D_{case} - D_{nom})/D_{nom}) * \Delta, \text{ where } \Delta = (V_{case} - V_{nom});$$
$$Kt = ((D_{case} - D_{nom})/D_{nom}) * \Delta, \text{ where } \Delta = (T_{case} - T_{nom});$$

The equation given above suffers from several problems when the
5 equation is used to calculate delays. First, if D_{nom} equals 0 or is very small, Kp, Kv
and Kt approaches infinity, thus, producing an invalid result. Second, the sign of D_{nom}
(whether positive or negative) affects the result. For example, if the voltage supplied
to a cell is reduced, the delay of the cell is suppose to increase, but if D_{nom} is a
negative number and the above equation is used to calculate D_{case} for a lower supply
10 voltage, the calculated delay, D_{case}, of the cell decreases instead of increasing, as is
the case in the real world. Thus, the sign of D_{nom} may affect the outcome and
produce an incorrect result. Third, the above equation is inaccurate because the
equation is based on data sampling at a single point and uses linear curve fitting
scheme to find the new delay. Fourth, the above equation is not suitable for derating
15 setup and hold times. Finally, the above equation does not capture the dependancy
of the delay on the fanout and the input ramptime (defined as the time duration an
input signal takes to switch between two logic levels completely) of the cells.

SUMMARY OF THE INVENTION

20 It is an object of the present invention to provide methods for
calculating delays for cells in an ASIC, which obviate for practical purposes the
above mentioned limitations.

According to an embodiment of the present invention, the delays,
including the propagation delays and the setup/hold time delays, are computed by
25 considering not only the process (P), voltage (V), temperature (T) but also input
ramptime (R) and output load or fanout (F) of the cells by fitting the delay at four
corner points for derated PVT condition into a non-linear equation which is a function
of P, V, T, R and F. Thus, in embodiments of the present invention, the delay
characterization is a five dimensional characterization process, and this
30 characterization space is split into (P,V,T) characterization and (R,T) characterization
to reduce the characterization time and resources. The present invention provides
for accurate calculation of delays for cells in ASICs.

6

Other features and advantages of the invention will become apparent from the following detailed description, taken in conjunction with the accompanying drawings which illustrate, by way of example, various features of embodiments of the invention.

5

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a simplified illustration of an integrated circuit chip on semiconductor material.

Fig. 2 is a flow chart outlining the method for computing delays of ASIC cells

10   in accordance with embodiments of the present invention.

Fig. 3 is a block diagram of a general-purpose computer system, representing one suitable computer platform for implementing the methods of the invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

15   In preferred embodiments of the present invention, delays are computed by considering not only the process (P), voltage (V), temperature (T) but also input ramptime (R) and output load or fanout (F) of the cells. Thus, in embodiments of the present invention, the delay is a five dimensional characterization, and the characterization is split into (P,V,T) characterization and

20   (R,F) characterization to reduce the characterization time and resources.

Fig. 2 illustrates a process for computing the delay of an ASIC cell in accordance with embodiments of the present invention. In step 100, data points for delays under the nominal condition (Dnom) for a cell are generated. The present invention generates the delay data points (Dnom) by conducting SPICE simulation

25   on the cell by using the parameters for the nominal condition (i.e., nominal P, V and T) and the transistor level netlist of the cell. However, during the simulation, the input ramptime (R) and the output load (F) of the cell are varied within a respective range. R is varied from Rmin (the minimum value for R) to Rmax (the maximum value for R), and F is varied from Fmin (the minimum value for F) to Fmax (the maximum

30   value for F) while P, V and T remain unchanged at their nominal values.

In certain embodiments of the present invention, approximately sixty values for Dnom are generated by varying the values of R and F. However, the following four values for Dnom (Dnom1, Dnom2, Dnom3 and Dnom4) are considered

7

to be most significant in calculating delays for the cell for the purpose of delay derating:

$Dnom1$ = the delay when R = Rmin, F = Fmin;

5     $Dnom2$ = the delay when R = Rmax, F = Fmin;

$Dnom3$ = the delay when R = Rmin, F = Fmax;

$Dnom4$ = the delay when R = Rmax, F = Fmax;

The values for P, V and T are set at : P = Pnom, V = Vnom, T = Tnom.

10   In step 110, the value for P is changed from Pnom to another type of process such as WNWP (weak N and weak P) process or SNSP (strong N or strong P) process while V and T remain unchanged. A SPICE simulation is conducted with the new P value. The R and F are varied as described above. The following new delays (Dp1, Dp2, Dp3 and Dp4) are generated with the SPICE simulation:

15

P = Pnew, Pnew ≠ Pnom, V = Vnom, T = Tnom;

$Dp1$ = the delay when R = Rmin, F = Fmin;

$Dp2$ = the delay when R = Rmax, F = Fmin;

$Dp3$ = the delay when R = Rmin, F = Fmax;

20     $Dp4$ = the delay when R = Rmax, F = Fmax.

In step 120, the value for T is changed from Tnom to another temperature value within the operating range of the cell while V and P remain at their nominal values. A SPICE simulation is conducted with the new T value. The R and

25   F are varied as described above. The following new delays (Dt1, Dt2, Dt3 and Dt4) are generated with the SPICE simulation:

P = Pnom, V = Vnom, T = Tnew, Tnew ≠ Tnom;

$Dt1$ = the delay when R = Rmin, F = Fmin;

30     $Dt2$ = the delay when R = Rmax, F = Fmin;

$Dt3$ = the delay when R = Rmin, F = Fmax;

$Dt4$ = the delay when R = Rmax, F = Fmax.

8

In step 130, the value for V is changed from Vnom to another voltage value within the operating range of the cell (e.g., 95% of Vnom or 105% of Vnom) while T and P remain at their nominal values. A SPICE simulation is conducted with the new V value. The R and F are varied as described above. The following new delays (Dv1, Dv2, Dv3 and Dv4) are generated with the SPICE simulation:

$P = Pnom, V = Vnew, Vnew \neq Vnom, T = Tnom;$

$Dv1 = $ the delay when $R = Rmin, F = Fmin;$

$Dv2 = $ the delay when $R = Rmax, F = Fmin;$

$Dv3 = $ the delay when $R = Rmin, F = Fmax;$

$Dv4 = $ the delay when $R = Rmax, F = Fmax.$

In step 140, the following equations are created for the process variation:

$Dp1 = Dnom1 + (m1p * R + m2p * F + Ap * R * F + Cp); R = Rmin, F = Fmin.$

$Dp2 = Dnom2 + (m1p * R + m2p * F + Ap * R * F + Cp); R = Rmax, F = Fmin.$

$Dp3 = Dnom3 + (m1p * R + m2p * F + Ap * R * F + Cp); R = Rmin, F = Fmax.$

$Dp4 = Dnom4 + (m1p * R + m2p * F + Ap * R * F + Cp); R = Rmax, F = Fmax.$

There are four unknowns in the above equations: m1p, m2p, Ap and Cp. The four unknowns are coefficients. Since there are four unknowns with four equations, the values for m1p, m2p, Ap and Cp can be solved.

Similarly, the following four more equations are created for the temperature variation:

$Dt1 = Dnom1 + (m1t * R + m2t * F + At * R * F + Ct); R = Rmin, F = Fmin.$

$Dt2 = Dnom2 + (m1t * R + m2t * F + At * R * F + Ct); R = Rmax, F = Fmin.$

$Dt3 = Dnom3 + (m1t * R + m2t * F + At * R * F + Ct); R = Rmin, F = Fmax.$

$Dt4 = Dnom4 + (m1t * R + m2t * F + At * R * F + Ct); R = Rmax, F = Fmax.$

There are four unknowns in the above equations: m1t, m2t, At and Ct. The four unknowns are coefficients. Since there are four unknowns with four equations, the values for m1t, m2t, At and Ct can be solved.

9

Similarly, the following four more equations are created for the voltage variation:

$$Dv1 = Dnom1 + (m1v * R + m2v * F + Av * R * F + Cv); R = Rmin, F = Fmin.$$
$$Dv2 = Dnom2 + (m1v * R + m2v * F + Av * R * F + Cv); R = Rmax, F = Fmin.$$
$$Dv3 = Dnom3 + (m1v * R + m2v * F + Av * R * F + Cv); R = Rmin, F = Fmax.$$
$$Dv4 = Dnom4 + (m1v * R + m2v * F + Av * R * F + Cv); R = Rmax, F = Fmax.$$

There are four unknowns in the above equations: m1v, m2v, Av and Cv. The four unknowns are coefficients. Since there are four unknowns with four equations, the values for m1v, m2v, Av and Cv can be solved.

In step 150, after solving for m1p, m2p, Ap, Cp, m1t, m2t, At, Ct, m1v, m2v, Av and Cv, the coefficients are applied to the following equation to solve for any new delays for the cell:

$$Dnew = Dnom + (m1p * R + m2p * F + Ap * R * F + Cp) + (m1v * R +$$
$$m2v * F + Av * R * F + Cv) * (Vnew - Vnom) + (m1t * R + m2t * F + At$$
$$* R * F + Ct) * (Tnew - Tnom).$$

A new delay can be solved for any given new P, V, T, R and F by using the above equation. The value for Dnom may be retrieved from a table which has various values for Dnom at various R and F.

Generally, the methods described herein with respect to IC design will be practiced with a general purpose computer, either with a single processor or multiple processors. The methods described herein will also be generally implemented in an ECAD system running on a general purpose computer. Figure 3 is block diagram of a general purpose computer system, representing one of many suitable computer platforms for implementing the methods described above. Figure 3 shows a general purpose computer system 151 in accordance with the present invention. As shown in Figure 3, computer system 151 includes a central processing unit (CPU) 152, read-only memory (ROM) 154, random access memory (RAM) 156,

10

expansion RAM 158, input/output (I/O) circuitry 160, display assembly 162, input device 164, and expansion bus 166. Computer system 151 may also optionally include a mass storage unit 168 such as a disk drive unit or nonvolatile memory such as flash memory and a real-time clock 170.

5   CPU 152 is coupled to ROM 154 by a data bus 172, control bus 174, and address bus 176. ROM 154 contains the basic operating system for the computer system 151. CPU 152 is also connected to RAM 156 by busses 172, 174, and 176. Expansion RAM 158 is optionally coupled to RAM 156 for use by CPU 152. CPU 152 is also coupled to the I/O circuitry 160 by data bus 172, control bus 10   174, and address bus 176 to permit data transfers with peripheral devices.

I/O circuitry 160 typically includes a number of latches, registers and direct memory access (DMA) controllers. The purpose of I/O circuitry 160 is to provide an interface between CPU 152 and such peripheral devices as display assembly 162, input device 164, and mass storage 168.

15   Display assembly 162 of computer system 151 is an output device coupled to I/O circuitry 160 by a data bus 178. Display assembly 162 receives data from I/O circuitry 160 via bus 178 and displays that data on a suitable screen.

The screen for display assembly 162 can be a device that uses a cathode-ray tube (CRT), liquid crystal display (LCD), or the like, of the types 20   commercially available from a variety of manufacturers. Input device 164 can be a keyboard, a mouse, a stylus working in cooperation with a position-sensing display, or the like. The aforementioned input devices are available from a variety of vendors and are well known in the art.

Some type of mass storage 168 is generally considered desirable. 25   However, mass storage 168 can be eliminated by providing a sufficient mount of RAM 156 and expansion RAM 158 to store user application programs and data. In that case, RAMs 156 and 158 can optionally be provided with a backup battery to prevent the loss of data even when computer system 151 is turned off. However, it is generally desirable to have some type of long term mass storage 168 such as a 30   commercially available hard disk drive, nonvolatile memory such as flash memory, battery backed RAM, PC-data cards, or the like.

11

A removable storage read/write device 169 may be coupled to I/O circuitry 160 to read from and to write to a removable storage media 171. Removable storage media 171 may represent, for example, a magnetic disk, a magnetic tape, an opto-magnetic disk, an optical disk, or the like. Instructions for
5   implementing the inventive method may be provided, in one embodiment, to a network via such a removable storage media.

In operation, information is input into the computer system 151 by typing on a keyboard, manipulating a mouse or trackball, or "writing" on a tablet or on position-sensing screen of display assembly 162. CPU 152 then processes the
10   data under control of an operating system and an application program, such as a program to perform steps of the inventive method described above, stored in ROM 154 and/or RAM 116. CPU 152 then typically produces data which is output to the display assembly 162 to produce appropriate images on its screen.

Expansion bus 166 is coupled to data bus 172, control bus 174, and
15   address bus 176. Expansion bus 166 provides extra ports to couple devices such as network interface circuits, modems, display switches, microphones, speakers, etc. to CPU 152. Network communication is accomplished through the network interface circuit and an appropriate network.

Suitable computers for use in implementing the present invention may
20   be obtained from various vendors. Various computers, however, may be used depending upon the size and complexity of the OPC tasks. Suitable computers include mainframe computers, multiprocessor computers, workstations or personal computers. In addition, although a general purpose computer system has been described above, a special-purpose computer may also be used.

25   It should be understood that the present invention also relates to machine readable media on which are stored program instructions for performing the methods of this invention. Such media includes, by way of example, magnetic disks, magnetic tape, optically readable media such as CD ROMs, semiconductor memory such as PCMCIA cards, etc. In each case, the medium may take the form of a
30   portable item such as a small disk, diskette, cassette, etc., or it may take the form of a relatively larger or immobile item such as a hard disk drive or RAM provided in a computer.

12

Although the present invention has been described in detail with regarding the exemplary embodiments and drawings thereof, it should be apparent to those skilled in the art that various adaptations and modifications of the present invention may be accomplished without departing from the spirit and scope of the

5   invention. Accordingly, the invention is not limited to the precise embodiment shown in the drawings and described in detail hereinabove.

0182595.1